Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 70 (2015) 499 – 507

Procedia Computer Science

4th International Conference on Eco-friendly Computing and Communication Systems, ICECCS 2015

FR: A Recommender for Finding Faculty Based On CF Technique

Diksha Nagpal[a], Sumit Kaur[a], Shruti Gujral[a], Amritpal Singh[b]

[a]*Department of Computer Science,Chandigarh University,Gharuan Mohali and 140413, India*
[b]*Department of Computer Science,Thapar University,Patiala and 147001, India*

Abstract

Number of options available for online users for buying items is growing manifold; tool which has come to the rescue of online users is Recommender Systems. Recommender System is expanding as an influential research area and has its applications in various fields like e-commerce, business, entertainment, education, medical sciences and so on. The developed applications in education field are limited to recommend the articles, research papers, courses and books to students and research scholars. However, application of recommender system to recommend faculty of an education institution to the students, management and other members does not exist. So to further extend its usage in education domain, applications of recommender system to recommend faculty of education institution to the students, management and other members is proposed and evaluated in this work. The collaborative filtering (CF) is used to implement Faculty Recommender and comparison between the user-user CF and item-item CF results is done to find the optimal approach for this proposed Faculty Recommender system.

Peer-review under responsibility of the Organizing Committee of ICECCS 2015

Keywords: Recommender system;collaborative filtering,;coefficients; similarity.

1. Introduction

A system which generates outputs i.e. individualized recommendations for users or guides the user with personalized recommendations of the items which are found useful or interesting in a large space of possible options, is defined as recommender systems. Diverse parts of technologies and approaches are formed by the recommender systems. In both the communities' research and commercial, where numbers of approaches are recommended (suggested) for providing different recommendations, these recommender systems have become popular. Recommendation system targets to provide the first-rate recommendations to the users. On the basis of the ratings given to the items liked by users or by utilizing the information of neighbor users or by using the contextual

Peer-review under responsibility of the Organizing Committee of ICECCS 2015
doi:10.1016/j.procs.2015.10.091

information provided by users, recommendations are made. Neighbor users are those which form the community of the active user. The suggestions are made to user either by providing characteristics of the product or by summarizing the opinion and critiques of the neighbor users. RS is growing as a dominant research area and it has applications in variety of fields like e-commerce (ebay.com, Amazon.com), entertainment (Launch.com, Moviefinder.com), education (Coursera.org, Docear.org), tourism (travelocity.com, ski-europe.com), medical sciences (Drugstore.com, Practo.com) and so on. Researchers are trying to expand the horizon of application of recommender systems and they are also working on overcoming the limitations like cold start, scalability and sparsity.

The recommendation approaches can be categorized on the basis of the process of generating recommendations. There are different types of approaches [10] of recommender systems that vary in terms of the knowledge used, the recommendation algorithm followed, how the recommendations are generated, assembled and presented to the user. These different approaches are illustrated in Figure 1.1:

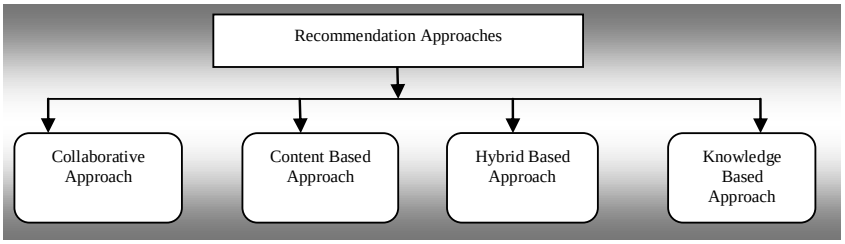

Fig. 1. Types of Approaches

The rest of paper is organised as follows: section II entails the technique used for developing Faculty Recommender (FR). Overview of working methodology of FR is provided in section III. Section IV presents and discusses the results to evaluate the better approach. Section V discusses the conclusion of the whole paper and the future scope.

2. Collaborative Filtering

Today there are number of collaborative filtering methods in service. These methods include different algorithms which compute predictions for preference of user and generate recommendations based on the preferences predicted. The task of recommendation can be thought of as the problem of information retrieval where the item domain and the profile of user preference are queried. Before studying the two basic approaches of Collaborative filtering we require to understand the concept of similarity.

2.1 Similarity

If two customers are found to be somewhat similar on the basis of some aspects and features of their profile or input data, then they are said to possess similarity between them. In the recommender system to find out the similarity between customers is a vital concept for generating recommendations. On the basis of their similarity the further steps are performed in generation of predictions. The similarity can be given as the correlation between customers. There are some measures used particularly for determining the correlation like Cosine, Pearson, Conditional probability etc. Some of these are discussed in below sections.

2.2 Approaches of Collaborative Filtering

2.2.1User–user collaborative filtering (UU-CF)

User–user collaborative filtering is one of the collaborative filtering based recommender systems. These algorithms are based on assumptions such as:

i. The taste of users will remain same in future
ii. Preferences of users remain consistent and stable

Items those are liked by users form the basis for recommendation. User-user Collaborative filtering was first used and introduced by the Group Lens [17] for Usenet article recommendation. User- user collaborative filtering is an uncomplicated algorithm that performs in steps; first step is to find the users which are similar to the current user on the base of the ratings in the past, the next step is to compute the recommendations and predict the ratings for the items that the current user will like. User-user Collaborative filtering method requires a matrix of the ratings, a similarity function to calculate the similarity and a method to predict the ratings by the use of ratings and similarities. The different methods which can be used to find out the similarity are the Pearson Coefficient, Cosine Coefficient and the Spearman rank correlation.

2.2.2 Item-item collaborative filtering (II-CF)

The UU-CF approach is quite effective but it suffers from the problem of scalability when the user data grows. In real time it becomes difficult to predict by scanning a huge number of neighbors. Item-item collaborative filtering was developed to extend the deployment of collaborative filtering. In real time II-CF [18, 19] allows to compute the predictions for large user databases. It uses the similarity between two items to compute the recommendations. The concept followed behind the item-item filtering is that by analyzing the information of the user about his purchasing history we can calculate the predictions that what the user would like in future. It is assumed that the user would purchase the items similar to those he had already purchased in past.

The coefficients like Cosine, Pearson, and Conditional Probability are used to measure the similarity between the items. The Cosine coefficient is analyzed to be best suitable for the item-item filtering as this measure outperforms with this filtering.

3. Methodology

For any educational institution, students are the central heart as all the uphill battle by faculty, management of the institution is done for ensuring students satisfaction. FR acts as an aid for students that facilitate them to steer the faculty dossier effortlessly. FR tries to understand the similarities among the faculty members which student has liked in the past (specific subject, expert area, languages, tools etc.). It computes and estimates the recommendation to the students of the faculty members. These recommendations could be helpful for students who are new to institute, it could help to assist them in finding and contacting faculty of their common area of interest. Similarly the management could be benefited by the information of faculty members for assigning any designated task as per their qualification and expertise. The development of this system would be beneficial for different institutional members. Figure 2 illustrates the architecture of the methodology followed for development and analysis of FR.

Following are the main steps performed for design and development of faculty recommender system.

1. Analysis of roles and responsibilities: Firstly for developing the Faculty Recommender, different roles of faculty members and their responsibilities towards the student and management is done. This analysis results in the output by enlisting numerous roles and responsibilities of faculty members in an institution.

2. Views of FRS: There are four different views of FRS, which includes: admin view, management view, faculty view and student view.

Admin View: Admin acts as a person with authorities to manage the whole database. He can create, update or delete any ID of the other members.

Management View: Management members can view the filtered information about their faculty members in the institution for their desired task.

Faculty View: Faculty members can filter the information about the other colleague members.

Student View: Students are the central heart of the whole system. They can view the filtered information about the faculty members and can give ratings to them.

3. Implementation of FRS: The implementation of FRS is done on .Net platform. Different views including main, admin, management, student and faculty view are developed using application language.

4. Creating and Updating Database: Using SQL server the complete structure of database is created. It has been created to manage the database of all users of the system; with help of these each user profile is created. Admin has the authorities to manage the whole database. The ratings provided by students to the faculty members are stored

in the database. Each time students' login for rating teacher database the updations in the database are done simultaneously.

5. Generating the ratings matrix from Database: The dataset of ratings which is stored with simultaneous students' activity is used for generating the rating matrix. This rating matrix is used in the process of generating recommendations both in user to user and item to item approach.

6. Generating Recommendations: Collaborative Filtering (CF) is used to develop the proposed recommender system. The CF involves two different approaches for generating recommendation. For FRS, both these approaches are applied to analyze the difference in **results.**

i. Generation of recommendations with User-User CF-Rating matrix is used as an input for this process. Firstly, Pearson Coefficient is computed for each student user, the similarity of each student user with other users is calculated to discover its neighbors who have high similarity. The generalized formula of Pearson Coefficient is used, which is given below:

$$sim(X,Y) = \frac{n\sum(XY) - \sum(X)\sum(Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}} \quad (1)$$

X and Y here represent the users between which the similarity is to be computed and n represents the total number of items for which the users provided ratings. After getting output of similarity rates, the filtering of the neighbors on some factors needs to be done. For this particular recommender system, a threshold value is set by getting the maximum value of the similarity and only considering the users who have similarity value 50% high than the maximum value for the particular student user.

Next is to predict the values of the ratings for those faculty members to whom the particular student user has not given any ratings. The prediction of the ratings is computed by formula shown below:

$$pred_{xi} = \frac{\sum sim_{xy}\, r_{yi}}{\sum |sim_{xy}|} \quad (2)$$

Here x and y represent the users, item is denoted by *i*.

To generate the recommendations the values of the prediction is used. Now, a threshold value of the prediction output is set. The predicted ratings which are greater than 3 for the faculty members on the 1-5 rating scale and also these faculty members are not previously rated by the particular student user, then the recommendation of these faculty members is generated and is shown on the view screen of student users.

The above mentioned steps for generating recommendations with User-User CF technique are proposed as an algorithm in Algorithm 1.

Algorithm: 1 UU - CF Technique for FR

Input: Set U= $\{u_1, u_2 \ldots\ldots u_n\}$ as set of users
Set I= $\{I_1, I_2 \ldots\ldots I_n\}$ as set of items
a and b are users of system, where a represents the active user and b is neighbor user.
$R_i^{u_a}$ is a function which refers to rating of ith item for u_a th user.

Begin:
1: Calculate values for α, β_1, β_2 where:

$$\alpha = \sum_{i=1}^{m}\left(R_i^{u_a}.R_i^{u_b}\right) - \left(\sum_{i=1}^{m} R_i^{u_a}\right)\left(\sum_{i=1}^{m} R_i^{u_b}\right)$$

$$\beta_1 = \sqrt{m\sum_{i}^{m}\left(R_i^{u_a}\right)^2 - \left(\sum_{i=1}^{m} R_i^{u_a}\right)^2}$$

$$\beta_2 = \sqrt{m\sum_{i}^{m}\left(R_i^{u_b}\right)^2 - \left(\sum_{i=1}^{m} R_i^{u_b}\right)^2}$$

2: Put the values calculated in step 1, in the below shown formula

$$S_{a,b} = \frac{\alpha}{\beta_1,\ \beta_2}$$

Here S represents the similarity between user a and b.
3: Set threshold T for setting the range of similar neighbors, here

$$T = 50\%\ of\ maximum\ similarity\ value$$

Users who have similarity value greater than 'T' are selected as neighbor users for computation of predicting rating values.
4: Calculate Prediction of ratings for user *a*, with use of output of from Step 3. Formula for predicting the rating value for an item *i* is:

$$pred = \frac{\sum P_{a,b} R_i^b}{\sum |P_{a,b}|}$$

5: Items which have predicted ratings greater than 3 are recommended to user *a*, i.e.

$$(pred > 3) \rightarrow recommended\ to\ user\ a$$

End

Algorithm: 2 II - CF Technique for FR

Input: Set U= $\{u_1, u_2 \ldots\ldots u_n\}$ as set of users
Set I= $\{I_1, I_2 \ldots\ldots I_n\}$ as set of items
$R_i^{u_a}$ is a function which refers to rating of i[th] item for u_a [th] user.

Begin:
1: Calculate values for α, β_1, β_2 where:

$$\alpha = \sum_{a}\left(R_i^{u_a}.\overrightarrow{Q_u^{u_a}}\right)\sum_{a}\left(R_j^{u_a}.\overrightarrow{Q_u^{u_a}}\right)$$

$$\beta_1 = \sqrt{\sum_{a}\left(R_i^{u_a}.\overrightarrow{Q_i^{i}}\right)^2}$$

$$\beta_2 = \sqrt{\sum_{a}\left(R_j^{u_a}.\overrightarrow{Q_j^{j}}\right)^2}$$

2: Put the values calculated in step 1, in the below shown formula

$$P_{a,b} = \frac{\alpha}{\beta_1,\ \beta_2}$$

3: Set threshold T for setting the range of similar neighbors, here

$$T = 50\%\ of\ maximum\ similarity\ value$$

Users who have similarity value greater than 'T' are selected as neighbor users for computation of predicting rating values.
4: Calculate Prediction of ratings for user *a*, with use of output of from Step 3. Formula for predicting the rating value for an item *i* is:

$$pred = \frac{\sum P_{a,b} R_i^b}{\sum |P_{a,b}|}$$

5: Items which have predicted ratings greater than 3 are recommended to user a, i.e.

$$(pred > 3) \rightarrow recommended\ to\ user\ a$$

End

The steps mentioned in section ii for generating recommendations with Item- Item CF technique are proposed as an algorithm in Algorithm 2.

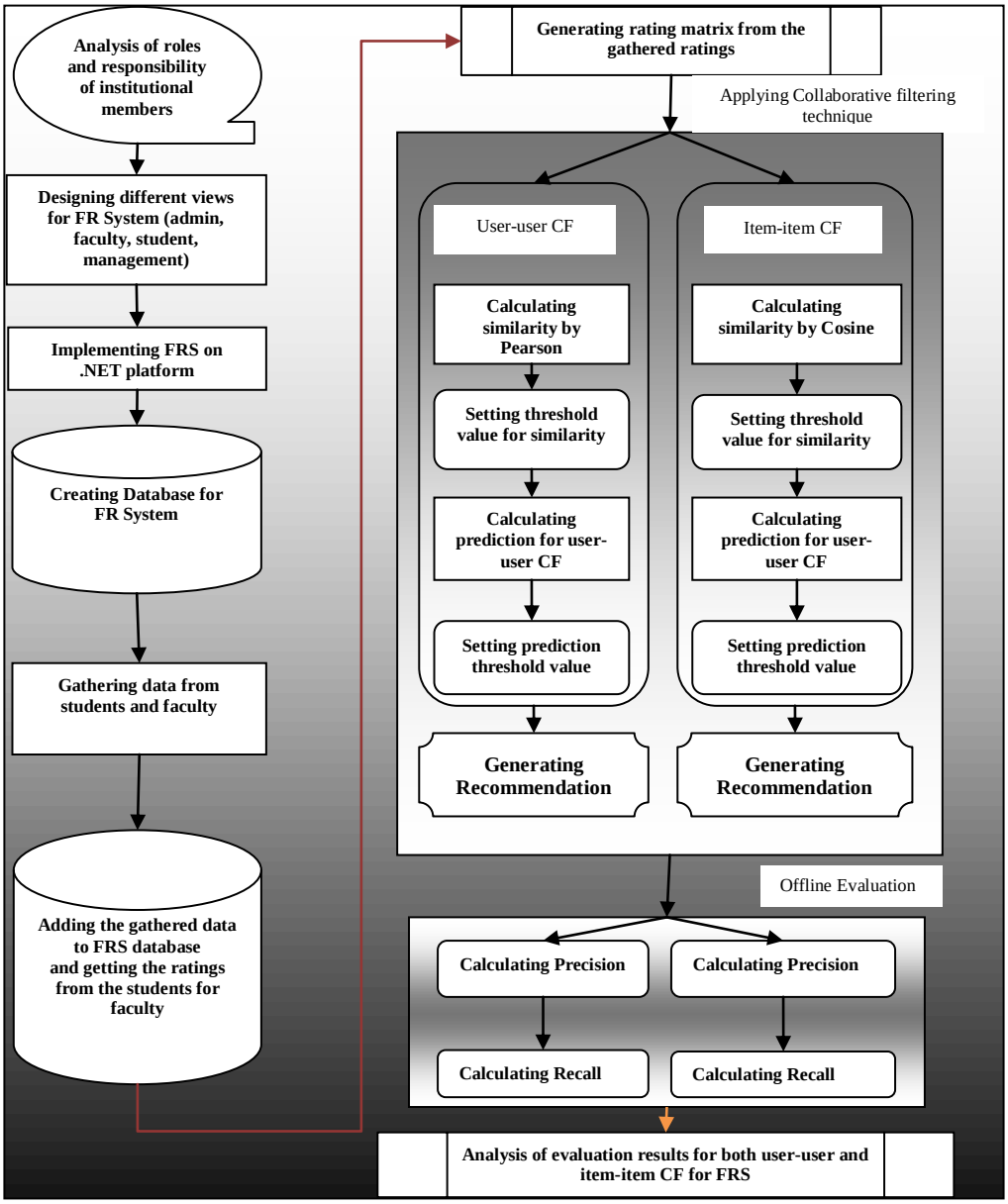

Fig 2. System Architecture

ii. Generation of recommendations with Item-Item CF-Rating matrix generated is also used in the item-item CF technique. The difference is that in this approach the recommendations are generated by focusing on the similarity

between the items which here represents the faculty members rather than the users. Cosine coefficient is used to compute the similarity between the faculty members from the database of each student user, the similarity of each faculty member with other faculty members is calculated to discover the members with high similarity.

$$sim(i,j) = \frac{\sum u \in U \left(r_{u,i} - \overline{r_u}\right)\left(r_{u,j} - \overline{r_u}\right)}{\sqrt{\sum u \in U \left(r_{i,p} - \overline{r_i}\right)^2}\sqrt{\sum u \in U \left(r_{u,j} - \overline{r_j}\right)^2}} \quad (3)$$

In above equation, $U = (u_1, u_2, \ldots, u_n)$ is the users set which rated both i and j items, R is n × m rating matrix and $\overline{r_u}$ denotes the average ratings.After getting output of similarity rates, the filtering of the items (faculty members) is done by setting the threshold. 50% of maximum value of similarity obtained is taken as the threshold value, only the values of the item that have similarity value higher than the threshold are selected.
The prediction of the ratings for the items for which student user has not given any ratings is computed by formula. The similarity results are used for predicting the ratings, as the results represent the similarity value between two items.

$$pred_{ui} = \frac{\sum sim_{ij} r_{uj}}{\sum |sim_{ij}|} \quad (4)$$

Here i and j represent the items, item is denoted by u.

The prediction results are used for generating recommendations. A threshold value of the prediction output is set like the **predicted ratings which are greater than 3** for the faculty members on the 1-5 rating scale and also these faculty members are not previously rated by the particular student user, then the **recommendation of these faculty members is generated** and is shown on the view screen of student users.

4. Results and Discussion

The proposed faculty recommender generated number of recommendations with both user-user and item-item filtering techniques. The student users were provided with these recommendations and asked to use the recommendations and rate the faculty members if they found any of them useful. Now, next step is to evaluate the performance of both the techniques. The below Table 1 represents the dataset of user and item which was collected by offline experiment and Table 2 shows the manually calculated recommendation results which were generated after applying algorithms 1 and 2:

Table 1. Dataset Collected By Offline Experiment

Total users (Students)	100
Total items(Faculty)	15

Table 2. Manual Calculated Results

Recommendations Using Algorithm 1		**Recommendations Using Algorithm 2**	
Total Recommendations generated	176	Total Recommendations generated	280
Recommendations used	115	Recommendations used	215

Calculating Precision and Recall: Recommendations generated from both the approaches are evaluated by using performance metrics. The whole experiment is an offline experiment and to evaluate the recommendation results confusion matrix is created to calculate precision and recall. With the help of Table 1 and 2, precision and recall values are computed for the User-User and Item-Item CF. The results are shown in tables 3 and 4 respectively.

Table 3. Precision and Recall For User-User CF

User-User CF	**Precision**	**Recall**
(Pearson Coefficient)	**58%**	**62%**

The precision value computed for User-User CF is 58% and recall value is 62%. Manual calculation is performed to calculate the values for confusion matrix and the values of confusion matrix are used for the calculation of performance metrics. The precision and recall value computed for Item-Item CF is 76% and 87% respectively.

Table 4. Precision and Recall For Item-Item CF

Item-Item CF (Cosine Coefficient)	Precision	Recall
	76%	87%

Discussion: Results computed by performance metrics shown in Table 3 and 4 are used to compare the User-User CF and Item-Item CF technique performance for the proposed Faculty Recommender.

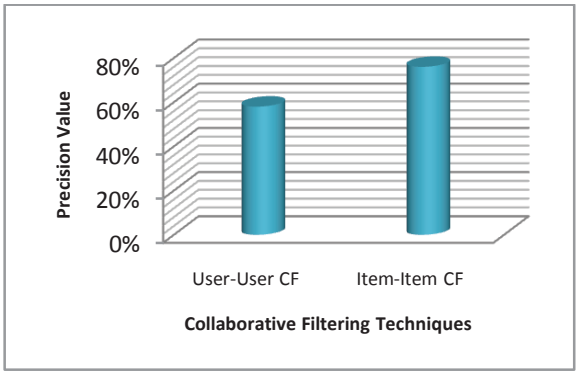

Fig 3. Graph Representation of Precision Results

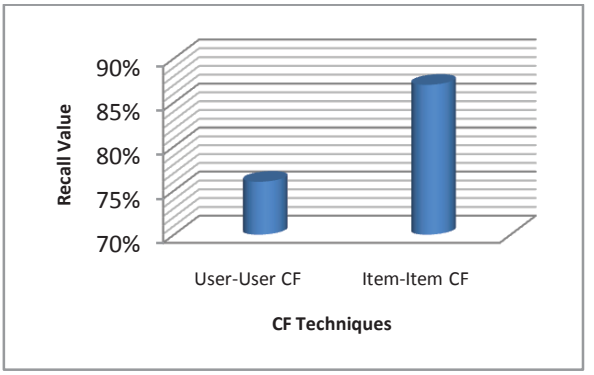

Fig 4. Graph Representation of Recall Results

From the above resulted graphs, it is clear that item-item CF performed better. Now, to evaluate that why the item-item CF technique is performing better for this proposed system, analysis is performed. During the analysis many points described the reason for the success of the item-item technique. The item-item filtering technique performed better for the proposed Faculty Recommender and both the results of the precision and recall are better than the user-user based technique. This situation might have occurred due to the below mentioned theories about the item-item technique.

i. The included dataset in the proposed recommender system have much greater number of users than the number of items i.e. student users are more than the faculty members. In [20, 21], it is mentioned that the more accurate results of recommendations are produced by the item-item filtering techniques in such type of cases.

ii. In the above mentioned situation, item-item filtering technique is preferred also because the computation time for computing similarity weights and the memory required is less in this technique.

iii. If the number of items is static as compared to the list of the users of the recommender system, than the item-tem technique is preferred. It is so, because the system is able to recommend items even to new users.

Therefore, these above mentioned reasons help to analyze the better working of the item-item CF technique for the Faculty Recommender.

5. Conclusion and Future Work

Primary focus of this work is to extend the usage of recommender systems in discipline of education. Recommender system for recommending the faculty of an institution to the students and management is designed and developed using collaborative filtering technique. Further, the development is done with two different approaches of collaborative filtering: user to user and item to item and in these techniques Pearson and Cosine coefficients are used to find the similarity among users respectively. The performance evaluation of these techniques for proposed Faculty Recommender system is done by finding the precision and recall values. This offline experiment is evaluated by performing manual operations. Precision and Recall values are computed for both the

techniques to compare their performance for this proposed system. Results of precision value for the both the techniques differ by 16%, item-item CF results in better results as compared to the other. Results of recall for the both the techniques differ by 25%, item-item CF results are better as compared to other. Item-Item CF technique performs better than the User-User CF technique as for this proposed faculty recommender it gave better results.

Future Work: The proposed recommender is a novel idea which can be extended further to improve the results. As collaborative filtering technique is used in this proposed faculty recommender, there are various other techniques which could be helpful for improving performance and results. Content based filtering, knowledge based or hybrid approach could be applied on the proposed recommender to increase the efficiency of faculty recommender. The dataset and the experiment performed is offline in this thesis work. The dataset was collected offline through a group of students who acted as users. This offline experiment could be converted to online experiment with use of online and real time dataset. This online transformation could help to calculate performance of faculty recommender more accurately.

References

1. Goldberg, David, David Nichols, Brian M. Oki, and Douglas Terry. "Using collaborative filtering to weave an information tapestry." *Communications of the ACM* 35, no. 12 (1992): 61-70.
2. Konstan, Joseph A., Bradley N. Miller, David Maltz, Jonathan L. Herlocker, Lee R. Gordon, and John Riedl. "GroupLens: applying collaborative filtering to Usenet news." *Communications of the ACM* 40, no. 3 (1997): 77-87.
3. "Short History of Collaborative Filtering" Moya K. Mason http://www.moyak.com/papers/collaborative-filtering.html.
4. Jannach, Dietmar, Markus Zanker, Alexander Felfernig, and Gerhard Friedrich. *Recommender systems: an introduction*. Cambridge University Press, 2010.
5. Shardanand, Upendra, and Pattie Maes. "Social information filtering: algorithms for automating "word of mouth"." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 210-217. ACM Press/Addison-Wesley Publishing Co., 1995.
6. Linden, Greg, Brent Smith, and Jeremy York. "Amazon.com recommendations: Item-to-item collaborative filtering." *Internet Computing, IEEE* 7, no. 1 (2003): 76-80.
7. Balabanović, Marko, and Yoav Shoham. "Fab: content-based, collaborative recommendation." *Communications of the ACM* 40, no. 3 (1997): 66-72.
8. Bennett, James, and Stan Lanning. "The netflix prize." In *Proceedings of KDD cup and workshop*, vol. 2007, p. 35. 2007.
9. Burke, Robin. "Hybrid recommender systems: Survey and experiments." *User modeling and user-adapted interaction* 12, no. 4 (2002): 331-370.
10. Ricci, Francesco, Lior Rokach, and Bracha Shapira. *Introduction to recommender systems handbook*. Springer US, 2011.
11. Sparling, E. Isaac, and Shilad Sen. "Rating: how difficult is it?." In *Proceedings of the fifth ACM conference on Recommender systems*, pp. 149-156. ACM, 2011.
12. Sarwar, Badrul, George Karypis, Joseph Konstan, and John Riedl. "Analysis of recommendation algorithms for e-commerce." In *Proceedings of the 2nd ACM conference on Electronic commerce*, pp. 158-167. ACM, 2000.
13. Schafer, J. Ben, Joseph A. Konstan, and John Riedl. "E-commerce recommendation applications." In *Applications of Data Mining to Electronic Commerce*, pp. 115-153. Springer US, 2001.
14. Adomavicius, Gediminas, and Alexander Tuzhilin. "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions." *Knowledge and Data Engineering, IEEE Transactions on* 17, no. 6 (2005): 734-749.
15. Herlocker, Jon, Joseph A. Konstan, and John Riedl. "An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms."*Information retrieval* 5, no. 4 (2002): 287-310.
16. Herlocker, Jonathan L., Joseph A. Konstan, Al Borchers, and John Riedl. "An algorithmic framework for performing collaborative filtering." In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 230-237. ACM, 1999.
17. Resnick, Paul, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. "GroupLens: an open architecture for collaborative filtering of netnews." In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pp. 175-186. ACM, 1994.
18. Deshpande, Mukund, and George Karypis. "Item-based top-n recommendation algorithms." *ACM Transactions on Information Systems (TOIS)* 22, no. 1 (2004): 143-177.
19. Sarwar, Badrul, George Karypis, Joseph Konstan, and John Riedl. "Item-based collaborative filtering recommendation algorithms." In *Proceedings of the 10th international conference on World Wide Web*, pp. 285-295. ACM, 2001.
20. Fouss, Francois, Alain Pirotte, Jean-Michel Renders, and Marco Saerens. "Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation." *Knowledge and data engineering, ieee transactions on* 19, no. 3 (2007): 355-369.